MULTIMEDIA UNIVERSITY

STUDENT ID NO

MULTIMEDIA UNIVERSITY

FINAL EXAMINATION

TRIMESTER 1, 2022/2023
TERM ID 2210

DMA5401 –INTRODUCTION TO PROBABILITY AND STATISTICS

(For Diploma students only)

13 FEBRUARY 2023
9.00 a.m - 11.00 a.m
(2 Hours)

INSTRUCTIONS TO STUDENT

1. This question paper consists of 5 pages (**3** pages with **4** questions and **2** pages for appendix)
2. Answer **ALL** questions. All necessary working steps must be shown.
3. Write all your answers in the answer booklet provided.
4. Statistical Tables will be provided.

QUESTION 1 [10 MARKS]
[CORRECT YOUR ANSWERS TO FOUR DECIMAL PLACES.]

(a) Hospital Cyberjaya keeps records of labor room (LR) traffic. Those records indicate that the number of pregnant women give births between 8.00 P.M. and 9.00P.M. has a Poisson distribution with parameter 6.9.

i. Determine the probability that, on a given day, the number of pregnant women who arrive at the labor room between 8.00 P.M. and 9.00 P.M. will be [3 marks]

i. Exactly three
ii. At most ten
iii. At least two
iv. Between four and six, inclusive

(b) From a sample of ten people surveyed, only 2% may consider on subscribing hardcopy magazine.

i. What is the probability that nobody would consider on subscribing hardcopy magazine? [1.5 marks]
ii. Find the probability that at least two people would consider on subscribing hardcopy magazine. [1.5 marks]

(c) The mathematics score of a class has a normal distribution with a mean of 75 and a standard deviation of 23.

i. Find the z-score if the student score is 63.5 [1 mark]
ii. If a student targeted to score between 80 and 72, calculate the percentage to achieve the target. [3 marks]

QUESTION 2 [10 MARKS]
[CORRECT YOUR ANSWERS TO FOUR DECIMAL PLACES.]

A. Ehlers et al. studied various characteristics of political prisoners From the Former East Germany and presented their findings in the paper "Posttraumatic Stress disorder (PTSD) following Political Imprisonment: The Role of Mental Defeat, Alienation, and perceived Permanent Change" (Journal of Abnormal Psychology, vol. 109, pp. 45_55). According to the article, the mean duration of imprisonment for 32 patients with chronic PTSD was 33.4 months. Assume $\sigma = 42$ months.

(i) Find the point estimator for μ. [1 mark]

(ii) Determine a 95% confidence interval for the mean duration of imprisonment, μ, of all East Germany Political prisoners with chronic PTSD. [3.5 marks]

(iii) Determine the margin of error, E in estimating μ. [1.5 marks]

(iv) Explain the meaning of E in this context in term of the accuracy of the estimate. [1 mark]

Continued...

(v) Find the sample size required to have a margin of error of 12 months and a 99% confidence interval. [3 marks]

QUESTION 3 [10 MARKS]
[CORRECT YOUR ANSWERS TO FOUR DECIMAL PLACES.]

(a) A common medicine for relieving serious pain is believed to be only 80% effective. A new medicine is used to a random sample of 100 adults who were suffering from serious pain and it shows that 85% received relief. Is this sufficient evidence to conclude that the new medicine is superior to the one commonly prescribed? Use a 0.05 level of significance? [4 marks]

(b) The means of the number of revolutions per minute (RPM) of two competing engines are to be compared. Thirty engines are randomly assigned to be tested. Both populations have normal distributions. Table below shows the result.

Engine	**Sample Mean of RPM**	**Population Standard Deviation**	**Number of engines**
1	1,500	50	30
2	1,600	60	30

i. State the null and the alternative hypothesis. [1 mark]

ii. Classify the hypothesis test as two tailed, left tailed, or right tailed. [0.5 mark]

iii. By using the hypothesis testing of two population mean, calculate the test statistic. [2.5 marks]

iv. Test whether the data indicate that Engine 2 has higher RPM than Engine 1? Test at a 5% level of significance [2 marks]

Continued...

QUESTION 4 [10 MARKS]
[FOR (b), CORRECT YOUR ANSWERS TO FOUR DECIMAL PLACES.]
It is believed that there is a linear relationship between house price (in thousand dollar), y, and its square feet, x. The following data show houses' square feet and their prices.

Square Feet, x	1100	1400	1425	1550	1600	1700	1725	1875	2350	2450
House Price (in thousand dollar), y	199	245	319	219	312	255	279	308	405	324

(a) Construct a table which contains x, y, x^2, y^2, xy and their individual sums. [2 marks]

(b) Find

i. the values of S_{XX}, S_{YY}, S_{XY}, y-intercept β_0 and slope β_1. [5 marks]

ii. the simple linear regression that fits the data obtained from (i). [0.5 mark]

iii. the house price when the house square feet is 2000. [0.5 mark]

(c) Compute and interpret the sample correlation coefficient, r. [2 marks]

Continued...

APPENDIX – KEY FORMULA

DISCRETE RANDOM VARIABLE, BINOMIAL DISTRIBUTION, AND POISSON DISTRIBUTION

- Standard deviation of a discrete random variable x : $\sigma = \sqrt{\sum x^2 P(X=x) - \mu^2}$
- Poisson probability formula : $P(X=x) = \dfrac{e^{-\lambda}\lambda^x}{x!}$

ESTIMATION

- The $(1-\alpha)100\%$ confidence interval for μ is

$\bar{x} \pm z_{\alpha/2}\sigma_{\bar{x}}$ where $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$

$\bar{x} \pm t_{\alpha/2,n-1}s_{\bar{x}}$ where $s_{\bar{x}} = \dfrac{s}{\sqrt{n}}$

- The $(1-\alpha)100\%$ confidence interval for p is

$\hat{p} \pm z_{\alpha/2}s_{\hat{p}}$ where $s_{\hat{p}} = \sqrt{\dfrac{\hat{p}\hat{q}}{n}}$

HYPOTHESIS TESTING

- **ONE population**

Mean:

$$Z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \quad \textbf{or} \quad t = \frac{\bar{x}-\mu}{s/\sqrt{n}}$$

Proportion:

$$Z = \frac{\hat{p}-p}{\sqrt{\dfrac{pq}{n}}}$$

- **TWO population**

Mean:

$$Z = \frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{\sqrt{\dfrac{\sigma_1^{\,2}}{n_1}+\dfrac{\sigma_2^{\,2}}{n_2}}}$$

Continued...

Proportion:

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}_p(1-\hat{p}_p)}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \qquad \text{where} \qquad \hat{p}_p = \frac{x_1 + x_2}{n_1 + n_2}$$

REGRESSION

- For the **least square regression equation**, $\hat{y} = \beta_0 + \beta_1 x$

  where $\beta_1 = \frac{S_{xy}}{S_{xx}}$ and $\beta_0 = \frac{1}{n}\left[\sum y - \beta_1 \sum x\right]$

- **Correlation coefficient, r**

  $$r = \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}}$$

  where

  $$S_{xx} = \sum x^2 - \frac{(\sum x)^2}{n}, \qquad S_{yy} = \sum y^2 - \frac{(\sum y)^2}{n} \text{ and}$$

  $$S_{xy} = \sum xy - \frac{(\sum x)(\sum y)}{n}$$

- **Error of sum of squares, SSE**

  $$SSE = S_{yy} - \frac{(S_{xy})^2}{S_{xx}}$$

- **Standard error of the estimate, s_e**

  $$s_e = \sqrt{\frac{SSE}{n-2}}$$

- **Regression *t*-Test**

  $$t = \frac{b_1 - \beta_1}{s_b} \qquad \text{where } s_b = \frac{s_e}{\sqrt{S_{xx}}} \qquad \text{and } s_e = \sqrt{\frac{S_{yy} - b_1 S_{xy}}{n-2}}$$

- **Correlation *t*-Test**

  $$t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}}$$

End of Page.